

Crystal structure of a repeating superhelix motif in the clathrin triskelion leg

Joel A. Ybe¹, Frances M. Brodsky¹, Thomas N. Earnest³,
Robert J. Fletterick² and Peter K. Hwang²

¹The G. W. Hooper Foundation, Department of Microbiology and Immunology,

²Department of Biochemistry and Biophysics,

University of California, San Francisco, California 94143, USA

³Advanced Light Source, Ernest Orlando Lawrence Berkeley National Laboratory,
University of California, Berkeley, California 94720, USA

INTRODUCTION

Clathrin is a triskelion-shaped cytoplasmic protein that polymerizes into a spherical, polyhedral lattice on intracellular membranes (Fig.1a)[1]. This reversible self-assembly process is required in the formation and turnover of protein-coated vesicles that mediate protein sorting during endocytosis and intracellular organelle biogenesis. The clathrin triskelion is a trimer of three heavy chain subunits (1675 amino acids/subunit), each folding into a filamentous leg and capable of binding a single light chain subunit (Fig.1b). Polyhedron formation involves interactions over the entire filamentous portions of the clathrin triskelion leg (residues 331-1522). To understand the molecular basis of clathrin self-assembly and its regulation, we have determined the crystal structure of residues 1210-1516 of the clathrin heavy chain to 2.6Å resolution. This segment typifies the structure of the whole triskelion leg and, in addition, comprises an important assembly control region of clathrin.

STRUCTURE DETERMINATION

A 55kDa region of bovine clathrin heavy chain spanning residues 1074-1522 was expressed as a selenomethionine-substituted protein in *E. coli* with an N-terminal polyhistidine tag for purification. The structure was determined by multiwavelength anomalous dispersion (MAD). A 4-wavelength MAD data set was collected from a single crystal (tetragonal space group I4122, one molecule per asymmetric unit) of selenomethionine-substituted clathrin heavy chain fragment (aa1074-1522) on beamline 5.0.2. The crystal was maintained at 100 K using an Oxford Cryostream, and data were collected in 1° oscillations using a 188mm x 188mm

Quantum-4 CCD detector system. 26124 unique reflections were measured, representing 72% completeness in the 30-2.5Å resolution range. The merging R-factor was 5%. Experimental phases produced a high-quality electron density map that was improved by density modification and phase combination. A molecular model of residues 1210-1516 was built into this map and refined to an R-factor of 24% and a free R-factor of 27% at 2.6Å resolution. The N-terminal residues 1074-1209 and C-terminus 1517-1522 are not included in the structural model, due to the

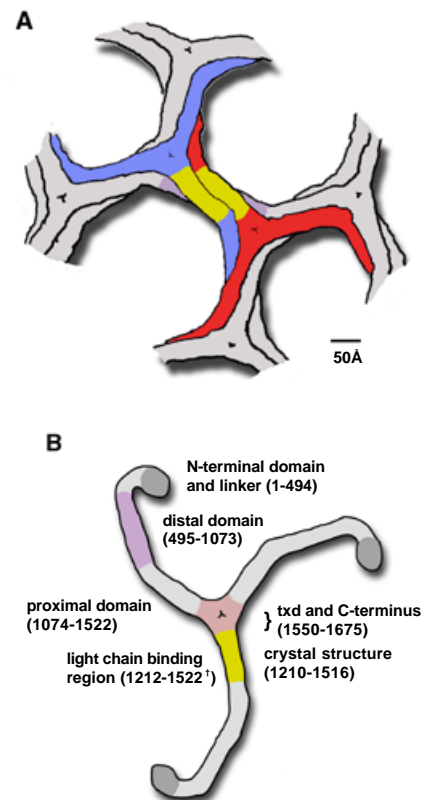


Figure 1. Organization of the clathrin polyhedron and triskelion. A, Portion of the clathrin polyhedron lattice [2], with a triskelion centered at each vertex. Two neighboring triskelions are colored red and blue. Yellow indicates portion of the proximal domain determined in crystal structure. B, Model of a single clathrin triskelion, formed by three heavy chains. Pink, TXD, the trimerization domain; dark gray, globular N-terminal domain.

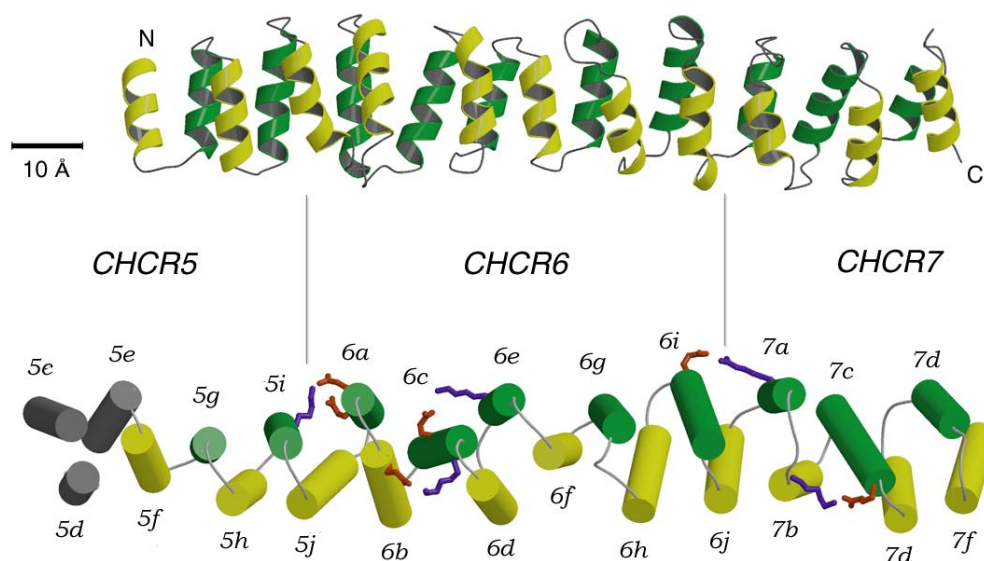


Figure 2. The three-dimensional structure of residues 1210-1516 in the clathrin heavy chain. Shown in two views related by 90° rotation about the length of the coil. Vertical segments demarcate the regions of clathrin repeats CHCR5, CHCR6 and CHCR7. Top, Ribbon diagram, viewed on the helix face. Yellow and green distinguish helices of separate faces. The “hairpin edge” is along the top of this view. Below, Cylinder representation, viewed on the hairpin edge, showing the relative orientation of individual helices, referred to as labeled. Gray cylinders indicate axes of helices 5c, 5d, and 5e, identified in the electron density map, but whose side chains could not be assigned. Red and blue side chains indicate acidic and basic salt bridge partners, respectively.

absence or poor quality of electron density in these regions of the map. Coordinates for the model have been submitted to the Brookhaven Protein Database (PDB code: 1b89).

A RIGHT-HANDED SUPERHELIX

The structural model of residues 1210-1516 in the clathrin heavy chain reveals a highly elongated structure composed entirely of short α -helices connected into a right-handed superhelix coil (Fig. 2). The α/α coiled fold is generated by stacking fundamental structural units of paired antiparallel α -helices joined by a hairpin turn. The canonical helical hairpin unit (helix-turn-helix-loop-) is 29 residues long, each helix of which averages 11 residues, or 3.5 turns. Ten complete helical hairpins (helices 5g-7f) plus helix 5f comprise the refined model. The electron density map identifies three more helices (5c, 5d, 5e), although this region of the map does not permit more detailed modeling.

Analysis of clathrin heavy chain sequences using the generalized profile method revealed a divergent but statistically significant repeat structure of the distal and proximal leg regions. Seven complete repeats (designated CHCR1-CHCR7) were found over the triskelion leg (residues 537-1566), suggesting a conservation of structural features throughout the leg. CHCR6 plus portions of CHCR5 and CHCR7 are represented in the solved crystal structure.

The organization of the helices within the repeat motif (*e.g.*, CHCR6) may be described as “2+4+4,” a single helical hairpin unit (*a* and *b*) followed by two sets of four-helix bundles (*cdef*, *ghij*). Almost all the conserved residues in the clathrin repeat are involved in helix

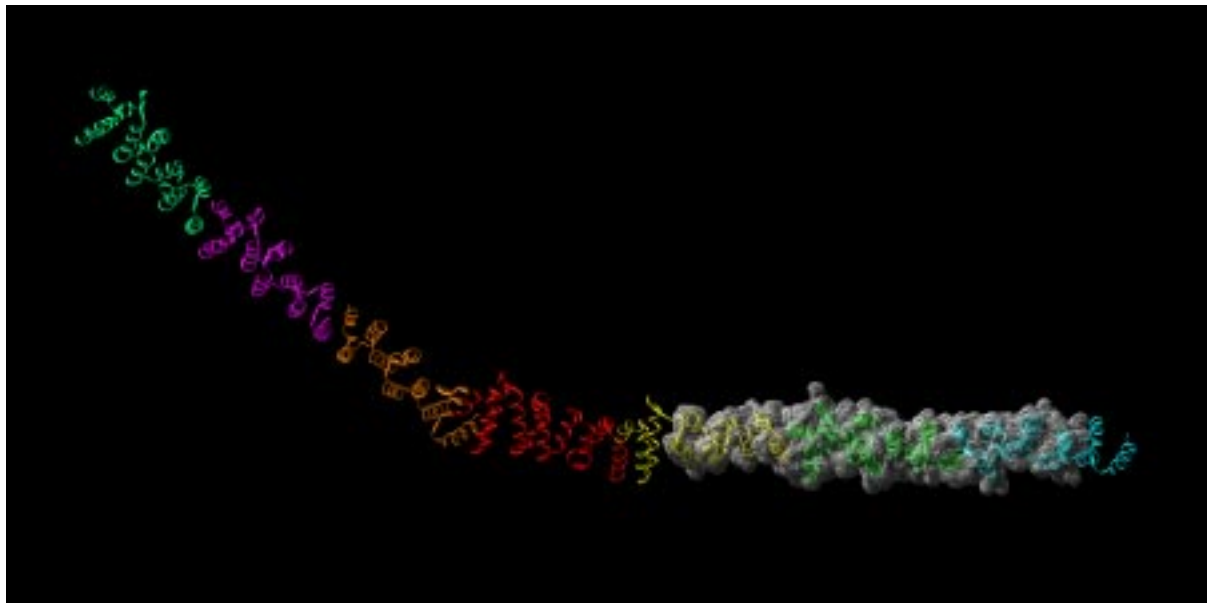


Figure 3. Ribbon structural model of the clathrin heavy chain leg domains, based on crystallographic identification of a repeated superhelix motif. Each of the seven repeat units is represented in a different color. The region of the solved crystal structure is indicated in space-filled representation.

packing and formation of the hydrophobic core, contributing to stability of the isolated filament. Furthermore, the distribution of conserved residues in helices *c*, *d*, *e*, and *f* suggests that the central helix bundle may be an invariant feature of the clathrin repeat.

Together the structural and sequence alignments predict that the entire filamentous portion of the clathrin triskelion leg is formed by a continuous superhelix comprised of clathrin repeats as defined here. The seven clathrin repeats making up the leg (residues 537-1566) predict that 70 helices fold into a 400Å long superhelix coil, stretching from the linker region at the N-terminal domain to the triskelion vertex (Fig.3). The uniform thickness and linear density of clathrin triskelions observed in electron micrographs [2] also suggest that the superhelix coiling of α -helices extends beyond the region of the seven aligned repeats and in fact may even provide the framework for trimerization interactions.

The distinguishing features of the clathrin superhelix (straightness and generally flat helix faces) likely have a functional impact on assembly. In particular, the superhelix may need to be straight with a slight twist in order to promote face-to-face interaction during self-assembly (Fig. 1a). In continuing studies, we hope to illuminate these and other features of regulated clathrin polymerization.

REFERENCES

1. Pearse, B. M. F. Clathrin and coated vesicles. *EMBO J.* **6**, 2507-2512 (1987).
2. Smith, C. J., Grigorieff, N. & Pearse, B.M.F. Clathrin coats at 21Å resolution: a cellular assembly designed to recycle multiple membrane receptors. *EMBO J.* **17**, 4943-4953 (1998).

This work was supported by the NIH and the U.S. Department of Energy Office of Basic Energy Sciences.

Principal investigator: Frances Brodsky, The G. W. Hooper Foundation, Department of Microbiology and Immunology, UCSF. Email: fmarbro@itsa.ucsf.edu. Telephone: 415-476-6405.